DOCUMENT RESUME

ED 203 576 EC 132 859

AUTHOR	Cooley, William W.: And Others
TITLE	Explaining Reading Performance of Learning Disabled Students.
INSTITUTION	Pittsburgh Univ., Pa. Learning Research and Development Center.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
PUB DATE	79
NOTE	22p.
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Curriculum: Expectation: *Learning Disabilities: *Models: Performance Factors: Primary Education: *Reading Achievement

ABSTRACT

The explanatory observational study investigated the relationships among curriculum content, instructional activity, teachers' expectations and reading performance for 52 children in six primary grade classrooms for the learning disabled. The analysis illustrates the utility of structural modeling in clarifying the probable causal relationships among the five variables studied. The model explains over 80% of the variance in end of year reading performance and reveals the ways in which the explanatory variables influence each other. Such models which may explain student achievement are seen as necessary prerequisites to evaluating educational innovations. (Author/CL)

EXPLAINING READING PERFORMANCE OF LEARNING DISABLED STUDENTS

William W. Cooley, Gaea Leinhardt, and Naomi Zigmond

Learning Research and Development Center

University of Pittsburgh

1979

The research reported herein was supported by the Learning Research and Development Center, sponsored in part by the National Institute of Education (NIE), United States Department of Health, Education, and Welfare. The opinions expressed do not necessarily reflect the position or policy of the sponsoring agency, and no official endorsement should be inferred. Joint authors listed alphabetically. The authors would like to acknowledge the important contributions to this effort made by the project staff: Doris Benson, Melanie Bowen, Jim DiCostanzo, Mary Ebert, Mary Engel, Connie Faddis, Bor-Jiin Mao, Andrea Seewald, and Ada Vallecorsa.

Abstract

This explanatory observational study investigated the relationships among curriculum content, instructional activity, teachers' expectations and reading performance for 52 children in six primary grade classrooms for the learning disabled. The analysis of these data illustrates the utility of structural modeling in clarifying the probable causal relationships among the five variables studied. The model explains over 80% of the variance in end-of-year reading performance and reveals the ways in which the explanatory variables influence each other.

EXPLAINING READING PERFORMANCE OF LEARNING DISABLED STUDENTS

William W. Cooley, Gaea Leinhardt, and Naomi Zigmond

Learning Research and Development Center
University of Pittsburgh

Over the past twenty years, there have been a number of attempts to identify significant classroom instructional factors that influence the development of reading skills in young children (e.g., Guthrie, Samuels, Martuza, Seifert, Tyler, & Edwall, 1976). These studies have not been particularly revealing for a variety of reasons, including unsuccessful efforts at implementing an experimental design, inadequate sampling of classroom events, and absence of a convincing causal model for guiding data collection and analysis. We hope to improve upon those previous efforts in a program of research that is currently underway at the Learning Research and Development Center. This paper reports some of the findings to date for one aspect of that work. Our purpose here is to explain, at a general level, the reading performance achieved by primary grade children classified as learning disabled and taught reading in six self-contained special education classrooms.

The specific educational environment in these learning disabilities (LD) classrooms has some unique features. In each classroom there is a teacher and an aide, and up to twelve children ranging in age from 6 to 11 years. The children are treated individually, both with respect to the rate at which they are taken through a sequence of instruction and the types of curricula to which they are exposed. Each child in an LD classroom has between two and ten streams of curricula for the learning of reading alone. In addition, each child can be on a unique time schedule. There is no single reading period for the whole class, but there are some periods in the day when it is likely that reading will take place for an individual child. Approximately two-thirds of each child's day is devoted to reading or language arts activities. These reading activities are spread throughout the day as are mathematics activities.

There are several reasons why we have chosen to study LD classrooms. We wanted settings in which there was a wide range of treatment variation, both among students within a classroom and among classrooms. It was also important to include students that represent a wide range of treatment variation, both among students within a classroom and among classrooms. It was also important to include students that represent a wide range of reading performance. These LD classrooms have that variation. In addition, we wanted classrooms with a small number of students per classroom so that the individual student could be the unit of observation. Since LD classrooms have a maximum of twelve students, they were ideal for developing this observational approach.

Study Design

An explanatory observational study of reading achievement was conducted in these learning disabilities classrooms. The study was explanatory as opposed to descriptive in that we proposed specific causal relationships among the variables studied. It was observational as opposed to experimental in that we examined educational experiences and outcomes as they occurred naturally. Thus, we measured reading achievement of students in the LD classrooms and examined relationships among variables representing teachers' expectations for student achievement, curriculum content, instructional activity, and final performance. The individual child was the unit of observation and analysis. Fifty-two learning disabled students comprised the sample.

In the simplest terms, we assumed that the end-of-year reading achievement of an individual student could be explained in terms of his/her initial achievement, the teacher's expectation about the student, the degree to which the content of the child's reading curriculum matched the content of the posttest, and the particular learning behaviors of the child between pre and posttest. In the most general form, this conceptualization can be expressed as:

Posttest = $\underline{f}$(pretest, teacher expectation, curriculum, instruction)

In the analysis of the data we report the results of a variety of analytical approaches. This is done for several reasons: (a) to describe the results in ways that have been generally employed in classroom research so that our findings can be easily compared to others; (b) to illustrate the advantages of structural modeling in non-experimental designs; and (c) to compare the results of a computer program for estimating linear structural relations (LISREL) with more familiar statistical procedures.

Pretest and Posttest

The reading subtest of the Comprehensive Test of Basic Skills (CTBS) (CTB/McGraw-Hill, 1974) was selected for pretest and posttest. Level A, B, or C was administered in the fall, depending upon the ability level of the child (determined by a previously administered Wide Range Achievement Test). Raw scores were converted to expanded standard scores in order to put all children on the same scale. The CTBS was readministered in the spring. Six months elapsed between pretest and posttest. Both pretest and posttest were administered individually by LRDC staff.

Teacher Expectation

Teacher expectation has been considered an important possible influence on students' achievement (Braun, 1976; Larsen, 1975; Rosenthal & Jacobson, 1968; Seaver, 1973). Therefore, the six teachers in this study were asked to rate each of their students in terms of expected end-of-year academic achievement. These ratings were obtained in February, using a seven-point rating scale. The rationale and statistical characteristics of this scale are reported by Benson (1979).

Curriculum Overlap

The curriculum overlap construct represents the relationship between curriculum content and the criterion measure. It is an estimate of how much of the criterion measure has been taught. Similar constructs have been used by other classroom researchers to explain student growth (Armbruster, Stevens, & Rosenshine, 1977; Cooley & Leinhardt, 1975, 1978; Fisher, Filby, Marliave, Cahen, Dishaw, Moore, & Berliner, 1978; Husen, 1967). There are two basic ways of estimating overlap, teacher judgment and student record. Teacher judgment is obtained at the time of posttesting by asking the teacher to mark, for each child, those items on the test for which the content has been taught (not necessarily learned). The overlap is the sum of all the items taught divided by the total number of items. The student record approach to curriculum overlap is obtained by creating a dictionary of all of the material (words and item forms) the student covers in all of the curricula. This curriculum dictionary is matched to a similar dictionary constructed for the test. If in a given test item a search of the student's curriculum dictionary indicates sufficient information to get the item right, the item is counted as covered. The overlap is the sum of the items judged covered by the curriculum divided by the total. (A rather complex system is used for paragraph estimates.) The advantage of the teacher estimate is that it is quick and easy to obtain and it includes instructional information not captured by student records. The disadvantage is that it may also include teacher bias. The advantage of the student record estimate is that it is obtained independent of teacher judgment. The disadvantage is that it underestimates what the child was taught on ditto worksheets, blackboard work, etc. For this study, we decided the teacher estimate of curriculum overlap was the better choice.

Observation of Reading Instruction

Most observations of classroom instruction focus on the teacher (Good & Brophy, 1971; Medley & Hill, undated). Those that include the student (Stallings, 1975; White, 1978) rarely focus on the content of the instructional activity. The exception to this is BTES (Fisher et al., 1978). However, the BTES system tends to reflect a combined content of curriculum and test presentation categories as opposed to an analysis of what the child needs to do in order to accomplish the task. The system used in this study was designed to assess what students actually do during reading instruction. A detailed analysis of reading behaviors coupled with an assessment of the tasks presented to children during reading formed the basis of the observational system. For example, during "seat work" a child may be copying words, silently reading, drawing lines between pictures and nouns, etc. The system was intended to be exhaustive, in the sense that all student behaviors during reading were classifiable and all of the student's time during observation could be accounted for.

At any given time a child being observed could be: reading aloud or silently, discussing (talking about or listening to) reading-related material, writing, waiting, off-task, engaged in some non-reading behavior such as mathematics or perceptual skills, out of the room or absent. If a child was reading, discussing, or writing, these activities could involve primarily letters, words, sentences, or paragraphs. Absent from the observation system as a separate category was something called "comprehension," because we do not view comprehension as separable from the specified activities. Also absent was modality of instruction (book, game, machine) because this is viewed as a non-instructional influence.

Observations were conducted by six observers and one trainer. Observations were made for 40, 80, or 120 consecutive minutes depending on the sampling plan. Each child was observed for 10 seconds. The observer then coded for 5 seconds and took 5 seconds to find the next student. After all children were observed once, the observer waited until a total of 5 minutes had elapsed since the start of the observation cycle. Then a second cycle was started. No cycle was started before the 5 minutes had elapsed. There were 8 cycles every 40 minutes.

Observers were trained with a self-training manual, then were field-trained in a seventh classroom (similar to the six in the study) until they reached criterion performance. After criterion was reached, the trainer monitored the observers with reliability checks every other week. Inter-observer reliabilities were also conducted at about the same rate. All observers observed pair-wise with all others, because having more than two observers in a classroom was too disruptive to teachers and students.

The stability or generalizability of the measures derived from the observation instrument, as estimated using intra-class correlation coefficients, was .68. Inter-observer reliability, estimated using intra-class correlations, was .90.

Four categories of activities coded on the observation instrument were considered to be reading events: reading aloud, reading silently, discussions of text, and writing. Reading aloud and reading silently are self-explanatory. Discussion of text involved any of the following child behaviors: listening to a story being read; listening to a summary, preparatory statement, review of the story (or sentence or paragraph) or questions about text; responding to questions about text; reciting summaries; etc. These listening or oral behaviors on the part of the child could precede or follow oral or silent reading of material. Writing involved activities such as: writing responses to questions; composing words, sentences, or paragraphs; rearranging letters or words; connecting or identifying words (or text), or pictures and text by lines or circles; copying; etc.

The four categories of reading activities accounted for approximately 12 minutes out of every 40 minutes of reading instruction. For this report, we decided to use silent reading and discussion to represent reading instruction. Writing was not included in the composite because, in these six learning disability classrooms, the majority of time spent in writing seems to have been spent in copying. This type of activity was not expected to relate to reading growth, and in fact did not. Oral reading was not included in the composite because the criterion measure did not have an oral reading component. Our measure of reading instruction is the time spent by children reading silently and discussing text. The stability of this particular measure, estimated using intra-class correlation coefficients, was .59.

Sampling Plan

The classroom observations of the 52 LD students took place over a 10-week period, from February 27 to May 12, 1978. The first step in drawing the sample of time to be observed was to obtain daily schedules from each of the six teachers. In general, each day was divided into eight 40-minute periods. Each period was classified as to the likelihood that reading or reading-related instruction would be going on during that period and the percentage of the students likely to be engaged in reading or reading-related activities. Since time schedules in these classrooms were highly individualized, the likelihood determinations had to be done at the student level. We then sampled randomly approximately 15% of the periods in which reading or reading-related activities could be expected for at least some of the students. Periods during which it was highly unlikely that any reading was going on for any student were not sampled at all. Each student was observed roughly 30 times over the 10 weeks. This number is approximate because in some classrooms reading instruction was concentrated in a few periods for all students at the same time, while in other classrooms, it was scheduled for small groups of students at a time and spread out over the entire day.

In order to obtain a single set of measures of reading instruction for each student, data from the 8 coded observations per period over the approximately 30 observations had to be combined. Our method was to sum all of the observations of a particular code and divide the sum by the number of 40-minute periods that the child was seen in school. This yielded an average amount of each code per 40 minutes of attended schooling. For each student, this average was then multiplied by the student's attendance rate. This yielded a weighted sum for each code for each child. The score represented time spent in each activity per 40-minute period.

Results and Discussion

Data from the five variables, pretest and posttest, teacher expectation, curriculum overlap, and reading instruction, were subjected to several levels of analysis. After means and standard deviations were derived for each variable, a correlation matrix was generated to determine the bivariate relationships. The strength of these intercorrelations suggested that more complex analyses should be carried out. These included investigating the relationships among variables when controlling for one variable at a time, and when controlling for several variables simultaneously. This done, we proceeded to a causal modeling of the variables.

Means and standard deviations for the five variables are in Table 1. The scale of the CTBS pretest and posttest is the expanded standard score. The pretest mean is roughly comparable to a 1.8 grade level. The grade equivalent for the posttest mean is approximately 2.1. The scale for teacher expectation is a seven-point rating scale. The mean is 4.27 with a 1.6 standard deviation, indicating that teacher responses were roughly in the middle of this particular scale.

Table 1

Means and Standard Deviations

(N = 52)

		Mean	Standard Deviations
1.	Pretest	273	48
2	Teacher Expectations	4.27	1.6
3.	Curriculum Overlap	59.1	33
4.	Reading Instruction	5.99	1.8
5.	Posttest	29[illegible]	49

Curriculum overlap is the teacher's estimate of the percentage of CTBS items that were taught to each student. The mean of 59.1% indicates that for the majority of children, teachers reported that over one-half of the items of the CTBS had been covered in reading instruction. However, the large standard deviation (33 percentage points) suggests that for some children, teachers estimated that very little of the test had been taught.

Reading instruction is scaled in terms of minutes of instruction out of a 40-minute period. The mean is 5.99 minutes with a 1.8 minute S.D., indicating that students spend approximately 6 of every 40-minute reading period in silent reading and discussion activities.

Correlation Analysis

Table 2 summarizes the correlations among the five variables. A review of the first column reveals that all variables have a substantial relationship with pretest. In addition to the expected high relationship between pretest and posttest (.82), the correlations show how dependent teacher expectation (.55), curriculum (.73) and instruction (.34) are upon initial abilities. One practice in correlational studies is to square the pre-post correlation ($.82^2 = .67$) and attribute exclusively to pretest the resulting portion of posttest variance explained.

Table 2
Correlation Matrix
(N = 52)

		1	2	3	4	5
1	Pretest	1.00				
2	Teacher Expectation	.55	1.00			
3	Curriculum Overlay	.73	.48	1.00		
4	Reading Instruction	.34	.42	.45	1.00	
5	Posttest	.82	.55	.81	.51	1.00
	Partial Correlations (with posttest controlling for pretest)		.42	.53	.43	1.00

The balance of the posttest variance is then left for the classroom processes and error to explain. Of course, a curriculum enthusiast might argue for an alternative interpretation, i.e., that two-thirds ($.81^2$) of the variance is explained by what the child was taught. Both arguments ignore the fact that these two highly correlated (.73) explanatory variables, pretest and curriculum, are inevitably confounded in individualized classrooms, and that it is not sensible to attribute to either one alone their joint effect.

The fifth row of the correlation matrix reveals that all of the variables also have a substantial relationship with posttest and that the magnitude of that relationship is the same as it is with pretest. These data suggest the need to "control" for pretest (using partial correlations) when describing the covariation of classroom process variables with posttest.

The last row of Table 2 gives the partial correlation of variables with posttest when pretest (or 67% of the posttest variation) is controlled. These partial correlations indicate the relationship between residuals from pretest for each predictor and posttest. The partials suggest that the pretest residuals for teacher expectation, curriculum overlap, and reading instruction are roughly comparable in their relationship to the posttest residuals.

One problem with partials is that they consider a five-variable network with only three variables at a time and assume that pretest is the only variable that needs to be controlled in considering a particular bivariate relationship. The evidence of high interrelationships among the classroom process variables indicates that procedures which consider all four predictors simultaneously are required.

Regression and Commonality Analyses

A multiple regression analysis of the correlation matrix in Table 2 yields the following standardized regression equation:

$$Z_5 = .40\, Z_1 + .20\, Z_2 + .36\, Z_3 + .13\, Z_4$$

The multiple correlation of .91 indicates that these four predictors explain about all of the reliable posttest variance. The coefficients in this regression evaluation suggest that pretest and curriculum are about equally useful in predicting posttest performance, with teacher expectation and reading instruction somewhat less useful.

A regression analysis of the observed variance-covariance matrix yields regression coefficients for raw scores as follows:

$$X_5 = .41\,X_1 + 6.3\,X_2 + .54\,X_3 + 3.4\,X_4 + 105.6$$

It is possible to interpret these raw score regressions as indicating, as in this example, that an increase of 1 minute in reading instruction (X_4) per 40 minutes would result in an increase in posttest of about 3.4 points (all other things being equal!). The problem with such an interpretation is that it assumes no measurement error or specification error, a rather untenable assumption.

Commonality analysis provides still another way of looking at these results. Table 3 summarizes the relative usefulness of these

Table 3

Commonality Results

Unique Effect	
1. Pretest	.066
2. Teacher Expectation	.027
3. Curriculum Overlap	.054
4. Instruction	.012
In-Common Effects*	
1, 2	.04
1, 3	.21
3, 4	.01
1, 2, 3	.16
1, 3, 4	.04
1, 2, 3, 4	.16
Multiple correlation = .91	

*Those > .01 are reported.

four predictors, both uniquely and in combination with other predictors. Notice that in terms of their unique contribution to the prediction of posttest, pretest, and curriculum are very similar. What the commonality results emphasize is the fact that the three large contributions to predicting posttest variance are the three in-common effects that include both pretest and curriculum.

Structural Modeling

To this point, we have described the relationships between posttest and the other variables taking two or three at a time, or simultaneously, but we have not postulated the way in which pretest, teacher expectation, curriculum overlap, and reading influence each other as they influence posttest. To do that, we need to build a structural model of the data that specifies the hypothesized relationships among variables in a causal network.

Figure 1 represents a possible structural model for these data. It suggests that posttest performance in reading is a function of pretest, the amount of silent reading and discussion a child engaged in (instruction), and the overlap between the content of the curriculum and the posttest. Teacher expectations are assumed to operate through curriculum overlap and reading instruction, but not affect posttest directly. Instruction and curriculum are both considered to be a function of the child's initial abilities and the teacher's expectations for the child. The amount of curriculum overlap is considered a function of initial abilities, teacher expectation, and the amount of instruction received.

The structural coefficients shown in Figure 1 reveal the relative influence of each variable on the others. In terms of explaining posttest, pretest and curriculum show major effects, and instruction a modest effect. These structural coefficients were estimated with the LISREL program (Joreskog & Sorbom, 1978). The program also provides a chi square test of how well the structural model seems to fit the data and clues as to what might be a better model. The chi square of 7.06 (_ndf_ = 1) indicates that differences between model and data are

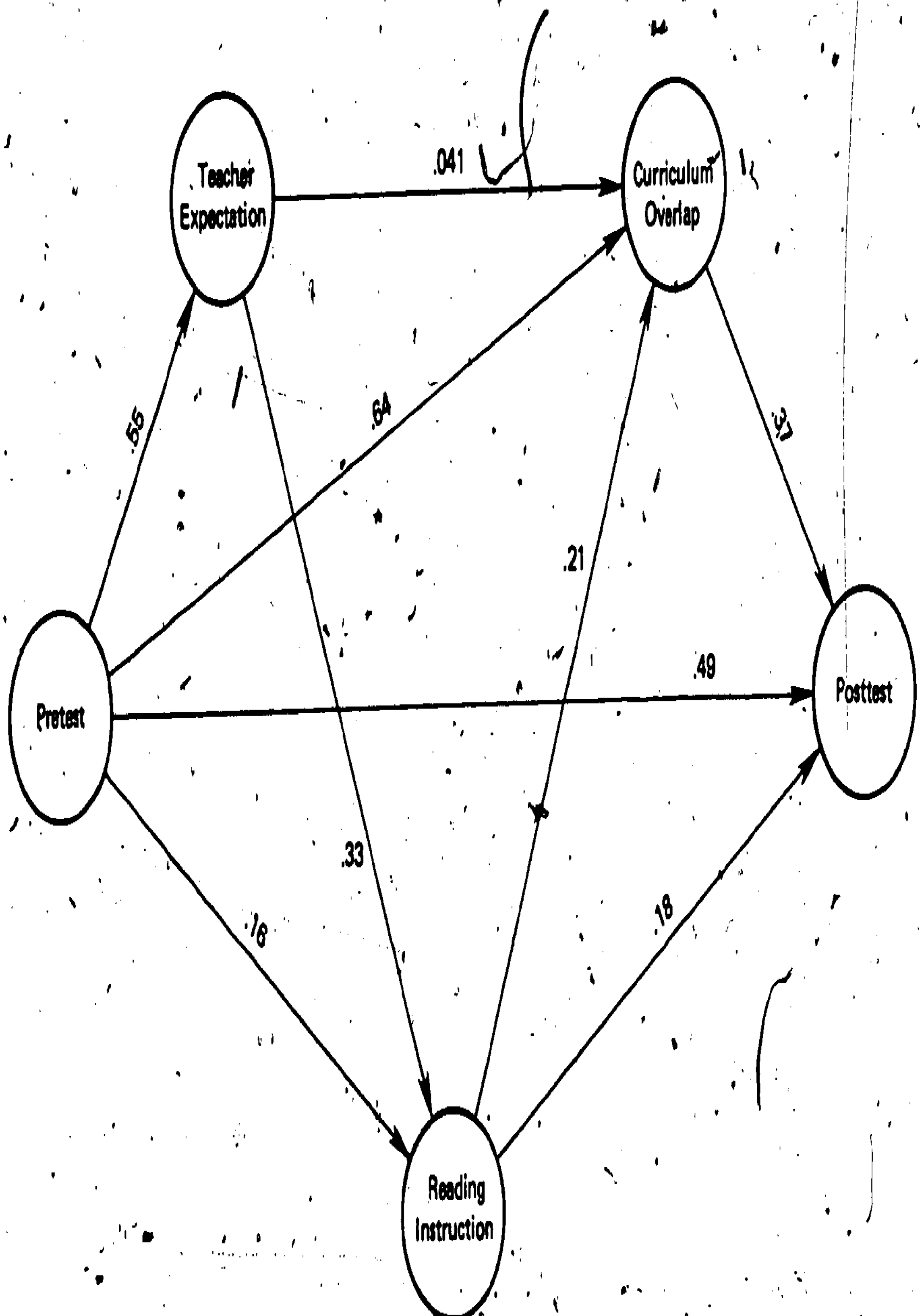

Figure 1. Hypothesized structrual model X^2 = 7.06
$\underline{p}$ = .008

probably not just due to chance. LISREL also computes the differences between the theoretical correlation matrix, as estimated from the structural model, and the observed correlations. These differences suggested that shifting the expectation effect from curriculum to posttest would produce a better fit between model and data.

Figure 2 illustrates a structural model in which the change was made. The chi square of .131 (ndf = 1) indicates a good fit of the data to the hypothesized model. It is useful to note that the structural coefficients in Figure 2 are identical to the standardized regression coefficients from the multiple regression. This occurred because we allowed all four explanatory constructs to show a direct influence on posttest. When that is done with only one indicator per construct, and when the distribution of the constructs is multivariate normal, ordinary least-squares regression yields the same structural coefficients as does LISREL's maximum likelihood procedures.

In the model of Figure 2, posttest is directly influenced by each of the four independent variables: pretest, teacher expectation, curriculum overlap, and reading instruction. The model also suggests interrelationships among the variables that help to explain the end-of-year reading performance achieved by learning disabled students.

Pretest not only influences posttest, but also has a relationship to teacher expectation, curriculum overlap, and, to a lesser extent, reading instruction. These results mirror the information presented in the correlation matrix, Table 2. They indicate that the classroom processes to which a student is exposed are at least in part a function of the student's initial performance level.

The model further suggests that teacher expectations, influenced by pretest, in turn influence the kind of reading instruction a child receives as well as his posttest performance. The apparent influence of expectation on instruction may represent a self-fulfilling influence at least to some degree. By self-fulfilling prophecy we refer to the classical description of teacher expectation (Rosenthal & Jacobson, 1968) in which an essentially unsubstantiated belief on the part of the teacher about a student results in modification of the teacher's

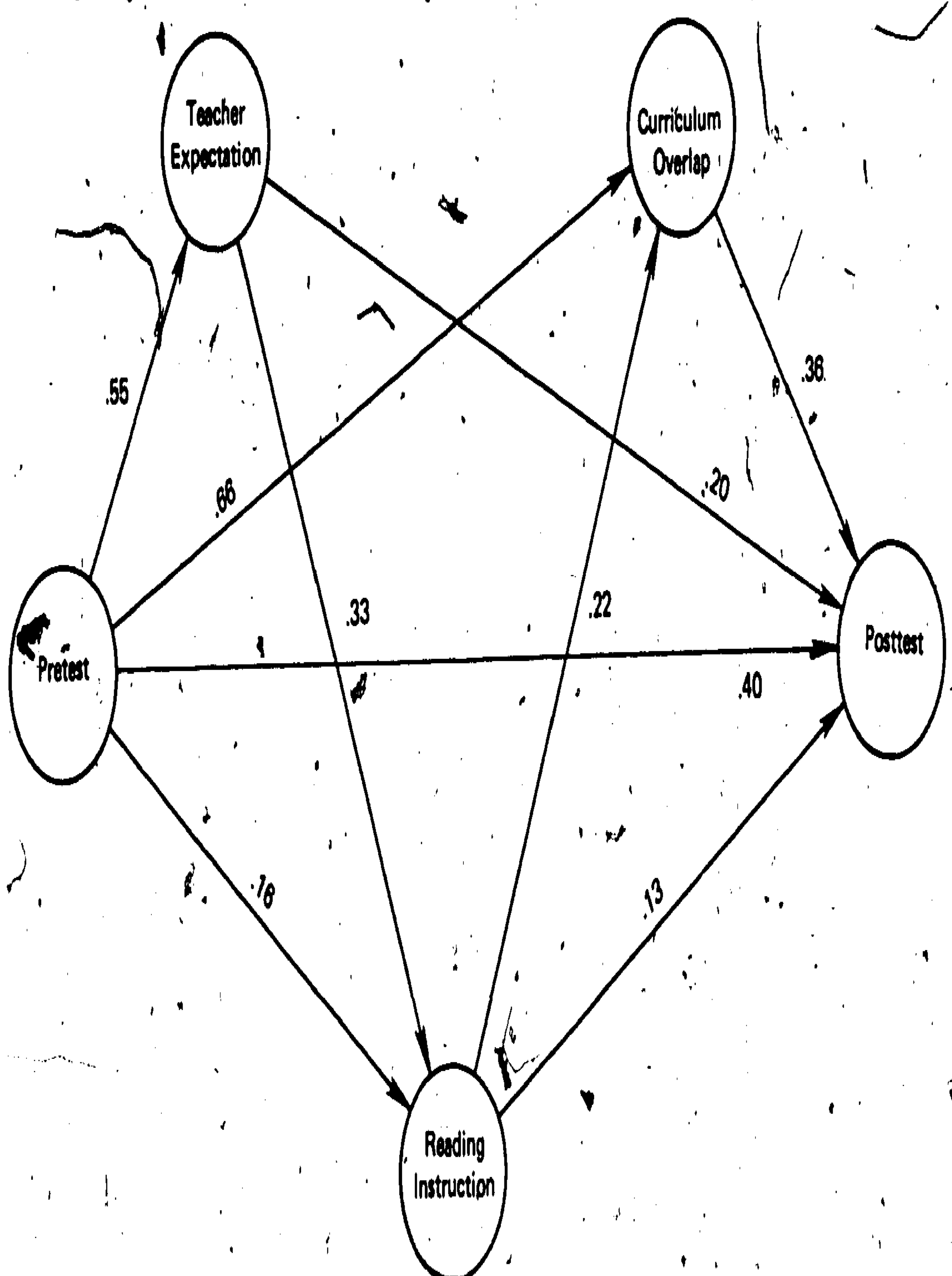

Figure 2. Structural model with a good fit to data $X^2 = .131$
$p = .72$

behavior toward a student that in turn affects student growth. However, the role of teacher expectation here seems to be more teacher prediction based on accurate knowledge than self-fulfilling prophecy. Some of the teachers have had as much as three years teaching the same students. And with the teacher expectancy rating obtained in February, even the new teachers may have had enough experience with their students to make the score simply an improvement over pretest as a predictor of posttest. The direct effect of expectations on posttest is either because teachers incorporate into their expectations some predictive information that was not captured by the pretest, or because the expectations were translated into behaviors that were not captured by the instruction or curriculum variables reported here.

The structural coefficients in Figure 2 suggest that curriculum overlap is in part a function of what the student already knows prior to instruction and the absolute quantity of reading instruction received during the course of the year. In turn, curriculum overlap strongly influences final test performance, i.e., children who were taught more of what the test measured did better on the test. This is not an astonishing finding. What is astonishing is the large number of studies that have attempted to show an effect for a particular approach to the teaching of reading that have not taken into account differences in overlap between the reading test and the reading approaches being contrasted. It may be that many studies of student achievement have attempted to get at subtle instructional differences and have tended to overlook the obvious.

Reading instruction, in the model elaborated in Figure 2, is seen to be only modestly influenced by initial performance and more substantially influenced by teacher expectations. Time spent in instruction is only slightly related to posttest but is significant in determining curriculum overlap, which in turn explains criterion performance. Through this latter mechanism, the influence of reading time appears quite relevant. The effectiveness of the instruction and curriculum variables are not dependent on pretest (i.e., no interactions were found). The more able students tended to cover more curriculum content and thus performed even better at posttest than one would have predicted from pretest alone.

Summary

The structural model presented in Figure 2 is a good fit to our current data. It represents, at a very general level, a causal network which explains the possible influences of classroom processes on the reading performance of primary grade children classified as LD and taught in self-contained special education classrooms. With the four variables--pretest, teacher expectation, curriculum overlap, and reading instruction--over 80% of the variance in criterion reading performance is explained.

During the 1978-79 school year, this study has been expanded to include 125 children from 11 LD classrooms. This represents the entire set of such classrooms in the Pittsburgh Public Schools. We have substituted the Diagnostic Reading Scales (Spache, 1972) for the CTBS so as to sample a wider range of initial and criterion reading behaviors. Also, we have modified our procedures for determining the sample of class time to be observed so that it will be possible to estimate the frequency with which reading or reading-related instructional events occur during the entire in-class experience of each child. In analyzing these new data, we also plan to utilize the full power of LISREL by employing multiple indicators of each construct. That should help to reduce the ambiguity that results from measurement error when estimating structural coefficients.

We feel that it is extremely important to work toward convincing causal models for explaining student achievement. As Cooley (1978) pointed out recently, such models are necessary prerequisites to the evaluation of educational innovations. Without a good understanding of the causal networks that are currently operating, it is essentially impossible to assess the impact of any manipulation of the variables in the network. We view this work as contributing toward that effort.

References

Armbruster, B. B., Stevens, R. J., & Rosenshine, B. Analyzing content coverage and emphasis: A study of three curricula and two tests (Technical Report #26). Urbana, Ill.: University of Illinois, 1977.

Benson, D. J. Teacher expectations and achievement of students with learning disabilities. Unpublished doctoral dissertation. Pittsburgh: University of Pittsburgh, Learning Research and Development Center, 1979.

Braun, C. Teacher expectation: Sociopsychological dynamics. Review of Educational Research, 1976, 46(2), 185-213.

Comprehensive tests of basic skills. Monterey, Calif.: CTB/McGraw-Hill, 1974.

Cooley, W. W. Explanatory observational studies. Educational Researcher, 1978, 7(9), 9-15.

Cooley, W. W., & Leinhardt, G. The application of a model for investigating classroom processes (LRDC Publication 1975/24). Pittsburgh: University of Pittsburgh, Learning Research and Development Center, 1975.

Cooley, W. W., & Leinhardt, G. The instructional dimensions study: The search for effective classroom processes (Final Report). Pittsburgh: University of Pittsburgh, Learning Research and Development Center, 1978.

Fisher, C. W., Filby, N. N., Marliave, R., Cahen, L. S., Dishaw, M. M., Moore, J. E., & Berliner, D. C. Teacher behaviors, academic learning time and student achievement: Final report of Phase III-B, Beginning Teacher Evaluation Study. Technical Report V-1. San Francisco, Calif.: Far West Laboratory for Educational Research and Development, 1978.

Good, T. L., & Brophy, J. E. Questioned equality for grade one boys and girls. Reading Teacher, 1971, 25, 247-252.

Guthrie, J. T., Samuels, S. J., Martuza, V., Seifert, M., Tyler, S. J., & Edwall, G. A study of the locus and nature of reading problems in the elementary school. Final Report. (NIE Contract No.: 400-75-0062). Washington, D.C.: National Institute of Education, 1976.

Husen, T., (Ed.) International study of achievement in mathematics: A comparison of twelve countries (Vol. II). New York: John Wiley & Sons, 1967.

Jöreskog, K. G., & Sörbom, D. Exploratory factor analysis program: A Fortran IV program. Uppsala, Sweden: University of Uppsala, Department of Statistics, 1978.

Larsen, S. C. The influence of teacher expectations on the school performance of handicapped children. Focus on Exceptional Children, 1975, 6(8), 1-13.

Medley, D. M., & Hill, R. A. Measurement properties of observation schedule and record, Form 4V, undated. (mimeo)

Rosenthal, R., & Jacobson, L. Pygmalion in the classroom. New York: Holt, Rinehart and Winston, 1968.

Seaver, W. B. Effects of naturally induced teacher expectancies. Journal of Personality and Social Psychology, 1973, 28, 333-342.

Spache, G. D. Diagnostic reading scales. Monterey, Calif.: CTB/McGraw-Hill, 1972.

Stallings, J. Implementation and child effects of teaching practices in Follow Through classrooms. Monographs of the Society for Research in Child Development, 1975, 40 (7-8, Serial No. 163), 1-119.

White, B. N. Manual for the analysis of behaviors in the classroom. Cambridge, Mass.: Research Institute for Educational Problems, Inc., 1978.